

Определение параметров функции развертки квадрупольного масс-спектрометра

¹А.С. Самохин*, ²Ю.А. Каламбет

¹Московский государственный университет имени М.В. Ломоносова,
Российская Федерация, 119991, г. Москва, Ленинские горы, 1-3

²ООО «Амперсенд», Российская Федерация, г. Москва, 123182, пл. Академика Курчатова, 2-2

*Адрес для переписки: Самохин Андрей Сергеевич, E-mail: andrey.s.samokhin@gmail.com

Поступила в редакцию 21 марта 2018 г., после исправления – 7 мая 2018 г.

Для решения большинства аналитических задач исследователю не нужно знать функцию развертки квадрупольного масс-спектрометра. В то же время все известные подходы, предназначенные для устранения искажений масс-спектров, зарегистрированных на склонах хроматографических пиков, основаны на априорном знании этой функции. Основная сложность заключается в вычислении параметра t_{overhead} , представляющего собой промежуток времени, необходимый прибору для подготовки к регистрации следующего масс-спектра. В настоящей работе мы разработали математический подход для решения этой задачи. Предлагаемый подход не требует наличия априорной информации о приборе и параметрах его работы. Подход основан на обработке фрагмента ГХ/МС данных, содержащего интенсивный отдельно стоящий хроматографический пик. Каждую масс-хроматограмму, содержащую пик аналита, обрабатывают независимо. Положение максимума пика (на каждой масс-хроматограмме) рассчитывают с точностью, превосходящей шаг дискретизации. Для этого данные аппроксимируют экспоненциально модифицированной функцией Гаусса. Показано, что положение максимума пика линейно зависит от значения m/z (что согласуется с теорией). Коэффициент наклона этой прямой используется для расчета параметра t_{overhead} . Показано, что воспроизводимость оценки параметра t_{overhead} главным образом определяется количеством аналита, введенного в прибор. Концентрация модельных соединений (полихлорированных бифенилов) составляла 0.3–10 нг мкл⁻¹ (пробу объемом 1 мкл вводили в режиме без деления потока). Когда концентрация модельных соединений составляла 3 нг мкл⁻¹ (или более), значение стандартного отклонения параметра t_{overhead} составляло 1.2 % от времени одного скана.

Ключевые слова: газовая хроматография/масс-спектрометрия, функция развертки, квадрупольный масс-анализатор, искажение масс-спектров.

For citation: *Analitika i kontrol'* [Analytics and Control], 2018, vol. 22, no. 2, pp. 168-176

DOI: 10.15826/analitika.2018.22.2.005

Estimation of quadrupole mass spectrometer scan function parameters

¹Andrey S. Samokhin*, ²Yury A. Kalambet

¹Lomonosov Moscow State University, Leninskie Gory 1-3, 119991 Moscow, Russian Federation

²Ampersand Ltd., pl. Academician Kurchatov 2-2, 123182 Moscow, Russian Federation

*Corresponding author: Andrey S. Samokhin, E-mail: andrey.s.samokhin@gmail.com

Submitted 21 March 2018, received in revised form 07 May 2018

In order to solve most analytical problems, the researcher does not need to know the scan function of a quadrupole mass spectrometer. However, at the same time, all known approaches for spectral skewing correction are based on a priori knowledge of this function. The main challenge is to calculate the overhead time, which is required to prepare the instrument for a new scan. In this paper we have developed a mathematical approach to solve this problem. The proposed approach does not require any prior information about the instrument and its operation parameters. The approach is based on the processing of GC/MS data fragment containing an intense baseline-separated chromatographic peak. Each extracted ion chromatogram

containing analyte signal is processed independently. The position of the peak maximum (in each extracted ion chromatogram) is calculated with an accuracy greater than one sampling interval. In order to do this, data are approximated by an exponentially modified Gaussian function. It is shown that the position of the peak maximum depends linearly on m/z value (this is consistent with the theory). The slope coefficient of this line is used to calculate the overhead time. It is shown that the reproducibility of the overhead time estimation depends mainly on the amount of analyte introduced into the instrument. Concentration of model compounds (polychlorinated biphenyls) is $0.3\text{--}10\text{ ng }\mu\text{L}^{-1}$ ($1\text{ }\mu\text{L}$ sample was injected in splitless mode). When the concentration of the model compounds is $3\text{ ng }\mu\text{L}^{-1}$ (or more), the standard deviation of the overhead time is 1.2 % of the total scan time.

Keywords: gas chromatography/mass spectrometry, scan function, quadrupole mass analyzer, spectral skewing.

Введение

В настоящее время метод газовой хроматографии/масс-спектрометрии (ГХ/МС) широко применяется в рутинном анализе. Наибольшее распространение получили ГХ/МС приборы, оснащенные одинарными квадрупольными масс-анализаторами (что в основном объясняется относительно невысокой стоимостью и компактными размерами таких приборов [1]). ГХ/МС приборы других типов применяют, как правило, в тех случаях, когда селективности масс-спектрометра с одинарным квадрупольным анализатором оказывается недостаточно. Анализ научных статей (опубликованных как в российских [2], так и в зарубежных журналах за последние 5–6 лет) показывает, что доля квадрупольных ГХ/МС приборов составляет порядка 70 %.

Если квадрупольный масс-анализатор работает в режиме регистрации полного масс-спектра, то в единицу времени через него могут пролетать ионы только с определенным значением m/z , задаваемым параметрами переменного электрического поля (создаваемого в пространстве между электродами). Чтобы зарегистрировать полный масс-спектр (или, другими словами, провести развертку масс-спектра), параметры поля меняют по определенной функции, обеспечивающей последовательную регистрацию ионов с различными значениями m/z [3–5]. Обычно развертку масс-спектра производят путем одновременного изменения величины постоянного напряжения (U) и амплитуды высокочастотного напряжения (V), сохраняя их отношение U/V постоянным. Под функцией развертки понимают зависимость изменяемых парамет-

тров поля (или регистрируемых значений m/z) от времени (рис. 1). При этом как напряжение, так и значения m/z регистрируемых ионов меняются во времени линейно (рис. 1). Следует отметить, что в квадрупольных масс-спектрометрах развертка может осуществляться путем изменения частоты высокочастотного напряжения, однако в этом случае зависимость изменяемого параметра от значения m/z будет нелинейной [3].

Для решения подавляющего большинства аналитических задач исследователю не нужно знать функцию развертки квадрупольного масс-спектрометра. Однако именно на априорном знании параметров этой функции основаны все известные подходы, предназначенные для устранения искажений масс-спектров, зарегистрированных на склонах хроматографических пиков [6–9]. Данные искажения масс-спектров характерны для всех сканирующих масс-спектрометров и наблюдаются в тех случаях, когда концентрация аналита в источнике ионов меняется в процессе развертки одного масс-спектра. Более подробная информация представлена, например, в работе [2].

Чтобы устранить искажения масс-спектров, зарегистрированных на склонах хроматографических пиков, необходимо знать диапазон сканирования, направление сканирования, а также значение параметра t_{overhead} (рис. 1), представляющего собой промежуток времени, необходимый квадрупольному масс-спектрометру для возвращения к исходным параметрам и подготовки к регистрации следующего масс-спектра. Диапазон и направление сканирования можно без труда извлечь из файла с ГХ/МС данными (даже если неизвестно, на каком приборе и при каких условиях были получены данные). С параметром t_{overhead} все обстоит сложнее. Если этот параметр неизвестен априори, то он, как правило, никак не учитывается. Такое упрощение можно использовать не всегда, поскольку значение параметра t_{overhead} может составлять десятки процентов от времени одного скана. В качестве примера на рис. 2 представлена зависимость параметра t_{overhead} от частоты сканирования для двух приборов, выпущенных фирмой Thermo Scientific.

Если параметр t_{overhead} не учитывается, то применение математического алгоритма для исправления исходных ГХ/МС данных приводит к инверсии наблюдаемых искажений. Если изначально на од-

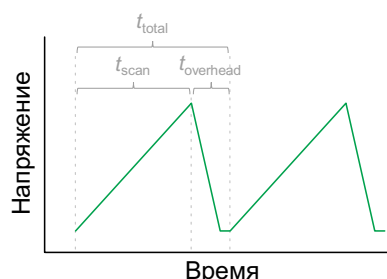


Рис. 1. Изменение напряжения (подаваемого на электроды) в процессе развертки масс-спектра

Fig. 1. Change in the voltage (applied to the rod electrodes) during mass spectrum scanning

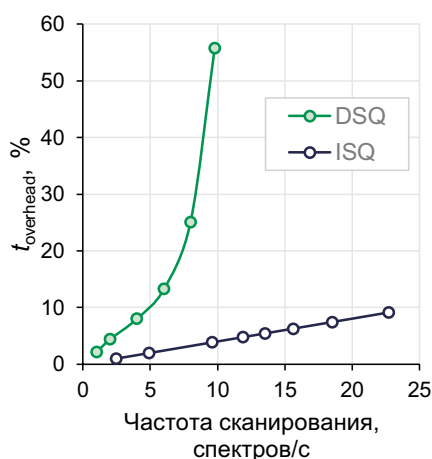


Рис. 2. Зависимость параметра t_{overhead} от частоты сканирования. Параметр t_{overhead} выражен в процентах от времени одного скана. Результаты получены для диапазона сканирования 50–550. DSQ и ISQ – одинарные квадрупольные масс-спектрометры, выпущенные фирмой Thermo Scientific

Fig. 2. Dependence of the overhead time on the acquisition speed. The overhead time is expressed as a percentage of the total scan time. The results are obtained for the scanning range of 50–550. DSQ and ISQ are single-quadrupole instruments manufactured by Thermo Scientific

ном из склонов хроматографического пика интенсивность масс-спектральных пиков, расположенных в области низких значений m/z , была занижена, то после применения математического алгоритма интенсивность этих пиков будет завышена, и наоборот. Это проиллюстрировано на рис. 3, на котором представлено, как изменяется отношение интенсивностей двух масс-спектральных пиков при движении по симметричному хроматографическому пику. Как видно из рис. 3, величина искажений масс-спектров, наблюдаемая после применения математического алгоритма, не учитывающего значение параметра t_{overhead} , тем больше, чем больше значение самого параметра t_{overhead} . Когда значение параметра t_{overhead} составляет 10, 20 или 30 % от времени одного скана, искажение масс-спектров после применения математического алгоритма составляет в среднем 9, 21 или 38 % от величины исходного искажения.

Наиболее просто извлечь значение параметра t_{overhead} из программного обеспечения, в котором задается режим работы масс-спектрометра. Именно так мы и поступили в нашей предыдущей работе [2], чтобы продемонстрировать возможности алгоритма, разработанного нами для устранения искажений масс-спектров, зарегистрированных на склонах хроматографических пиков. Однако значение параметра t_{overhead} может не отображаться в программном обеспечении (что, например, имеет место в случае масс-спектрометра «Agilent 5973», используемого в настоящей работе).

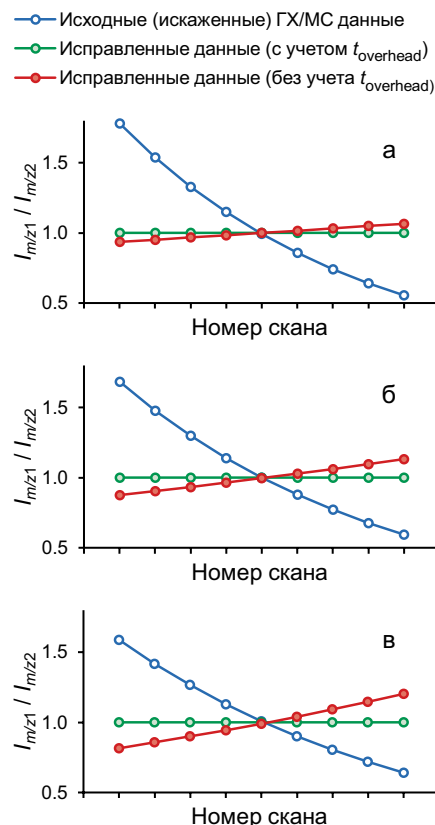


Рис. 3. Отношение интенсивностей двух масс-спектральных пиков ($I_{m/z1} / I_{m/z2}$), наблюдаемое в различных точках соответствующего хроматографического пика. Параметр t_{overhead} равен 10 (а), 20 (б) или 30 % (в) от времени одного скана. Отношение $I_{m/z1} / I_{m/z2}$ рассчитано только для тех точек хроматографического пика, интенсивность которых превышает 10 %

Fig. 3. Ratio of two mass spectral peaks ($I_{m/z1} / I_{m/z2}$) calculated for the different points of respective chromatographic peak. Overhead time is equal to 10 (a), 20 (б) or 30 % (в) of the total scan time. Ratio $I_{m/z1} / I_{m/z2}$ is calculated only for those points of the chromatographic peak which intensity exceeds 10 %

Цель настоящей работы заключалась в разработке математического подхода, предназначенного для расчета параметра t_{overhead} непосредственно из исходных ГХ/МС данных (без использования априорной информации о модели квадрупольного масс-спектрометра и параметрах его работы). Предлагаемый подход основан на том, что хроматографические пики одного и того же соединения, зарегистрированные на различных масс-хроматограммах, будут сдвинуты друг относительно друга на небольшую величину (не превышающую время развертки одного масс-спектра).

Экспериментальная часть

Общее описание подхода. В случае сканирующих приборов значения m/z , зарегистрированные в процессе развертки одного масс-спектра (то есть в пределах одного скана), фактически

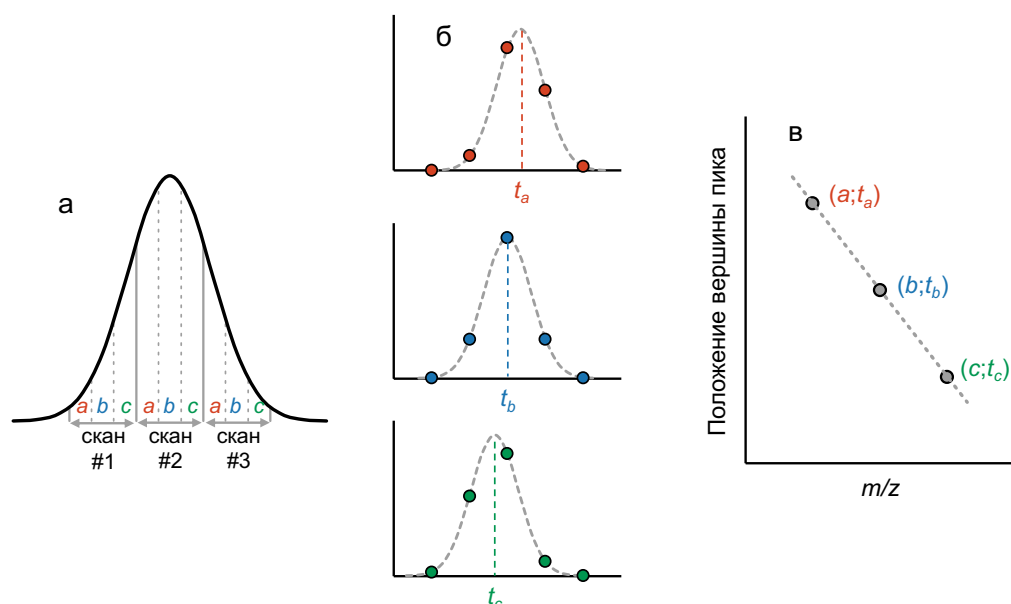


Рис. 4. Концентрационный профиль гипотетического соединения, элюируемого из хроматографической колонки (а). В масс-спектре соединения присутствуют три значения m/z (a, b и c). Масс-хроматограммы сдвинуты друг относительно друга на небольшую величину (б). Теоретически время удерживания (рассчитанное с точностью, превосходящей шаг дискретизации) должно линейно зависеть от соответствующего значения m/z (в)

Fig. 4. Concentration profile of a hypothetical compound eluted from a chromatographic column (a). Three m/z values (a, b and c) are present in the mass spectrum of the compound. Extracted ion chromatograms are shifted relative to each other by a small distance (б). Theoretically, retention time (calculated with an accuracy greater than one sampling interval) should depend linearly on the corresponding m/z value (в)

отвечают разным участкам непрерывного концентрационного профиля аналита (рис. 4а). Это приводит к тому, что масс-хроматограммы, представленные в координатах «Интенсивность – Номер скана», сдвинуты друг относительно друга на небольшую величину, не превышающую время развертки одного масс-спектра (рис. 4б). Набор дискретных точек определенной масс-хроматограммы можно использовать для восстановления формы непрерывного профиля пика. Положение максимума восстановленного пика зависит от значения m/z соответствующей масс-хроматограммы. Эта зависимость определяется функцией развертки и в случае квадрупольного прибора должна быть линейной (рис. 4в). Зависимость положения максимума пика от значения m/z может быть аппроксимирована прямой линией и использована для определения параметров функции развертки (рис. 4в). В этом случае знак углового коэффициента прямой (k) зависит от направления сканирования, а для расчета параметра t_{overhead} можно использовать следующую формулу:

$$t_{\text{overhead}} = 1 - |k| \cdot (mz_{\text{max}} - mz_{\text{min}} + 1),$$

где t_{overhead} – время (выраженное в долях от времени одного скана), необходимое прибору для подготовки к регистрации следующего масс-спектра; k – угловой коэффициент прямой (рис. 4в); mz_{min} и mz_{max} – значения m/z , отвечающие началу и концу диапазона сканирования соответственно.

Определение времени удерживания с точностью, превосходящей шаг дискретизации, проводили, интерполируя или аппроксимируя исходные данные рядом математических функций. В частности, нами были использованы следующие подходы:

- интерполяция параболой (по трем точкам);
- аппроксимация параболой (по пяти точкам);
- аппроксимация всей области хроматографического пика функцией Гаусса;
- аппроксимация всей области хроматографического пика экспоненциально модифицированной функцией Гаусса (ЭМГ) [10]. Следует отметить, что в этом случае строили две модели. При построении первой модели оптимизировали все четыре параметра функции (H , μ , σ и τ). При построении второй модели оптимизировали только два параметра (H и μ), поскольку два другие параметра (σ и τ), определяющие форму пика, были зафиксированы. При этом для всех масс-хроматограмм, на которых присутствовал пик аналита, использовали одни и те же параметры σ и τ (которые представляли собой медианные значения, рассчитанные для семи масс-хроматограмм, характеризующихся наибольшим отношением сигнал/шум).

Функция Гаусса описывается следующей формулой:

$$F_{\text{Gaussian}}(t) = H \cdot \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right),$$

где H – высота пика; μ – положение максимума пика; σ – параметр, задающий ширину пика (ширина пика на половине высоты: $W_{FWHM} = 2.35\sigma$).

ЭМГ описывается следующей формулой:

$$F_{EMG}(t) = H \cdot \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{t - \mu}{\tau}\right) \cdot \operatorname{erfc}\left(\frac{1}{\sqrt{2}} \cdot \left(\frac{\sigma}{\tau} - \frac{t - \mu}{\tau}\right)\right),$$

где H – параметр, задающий высоту пика (значение ординаты в максимуме не равно значению H); μ – параметр, задающий положение максимума пика; σ и τ – параметры, задающие ширину и степень асимметрии пика; $\operatorname{erfc}(x)$ – дополнительная функция ошибок. Следует подчеркнуть, что вычисление ЭМГ проводили согласно рекомендациям, представленным в работе [11] (используя как представленную выше, так и альтернативную формулу [12]).

Предварительная обработка данных. На первом этапе для каждого рассматриваемого соединения определяли набор значений m/z , для которых отношение сигнал/шум превышало 10 (в настоящей работе под шумом понимали четырехкратное стандартное отклонение шума). Затем проводили вычитание базовой линии, которую моделировали прямой линией. При этом все фрагменты масс-хроматограмм, содержащие пик интересующего соединения, обрабатывали независимо.

Вычисления. Аппроксимацию параболой осуществляли, используя классический метод наименьших квадратов. Функция Гаусса и ЭМГ являются нелинейными относительно оцениваемых параметров. В связи с этим аппроксимацию проводили с использованием алгоритма Левенберга–Марквардта. Коэффициенты аппроксимирующих функций вычисляли, используя программу консольной строки, написанную нами на языке C. Последующую обработку проводили в программном обеспечении Excel (Microsoft).

Условия ГХ/МС анализа. Смесь полихлорированных бифенилов анализировали на квадрупольном масс-спектрометре «Agilent 5973» (Agilent), соединенном с газовым хроматографом «Agilent 6890» (Agilent). Пробу вводили в режиме без деления потока (вентиль делителя потока открывали через 40 с). Разделение анализов проводили на капиллярной колонке Restek RTX-5 ms, 30 м × 0.25 мм × 0.25 мкм. Температура источника ионов – 200°C. Энергия ионизирующих электронов – 70 эВ. Частота сканирования – 4.4 спектр с⁻¹. Диапазон сканирования от 50 до 370. Направление сканирования – от максимального значения m/z к минимальному.

Модельная смесь содержала следующие соединения: 2-хлорбифенил, 2,3-дихлорбифенил, 2,4',5-трихлорбифенил, 2,2',4,4'-тетрахлорбифенил, 2,2',3',4,6-пентахлорбифенил, 2,2',4,4',5,6'-гексахлорбифенил, 2,2',3,3',4,4',6-гептахлорбифенил и 2,2',3,3',4,5',6,6'-октахлорбифенил. Гептахлорбифенил и октахлорбифенил не рассматривали в данной работе, поскольку в их масс-спектрах присутствуют значения m/z , превышающие верхнюю границу

диапазона сканирования. В работе использовали несколько модельных растворов с различными концентрациями анализов: 0.32, 1, 3.2 и 10 нг·мкл⁻¹. Если не указано особо, то концентрация модельных соединений составляла 1 нг·мкл⁻¹. Каждый модельный раствор был проанализирован три раза.

Результаты и их обсуждение

В настоящей работе в качестве модельных соединений использовали полихлорированные бифенилы. Выбор этого класса соединений объясняется рядом причин. Во-первых, в масс-спектрах этих соединений присутствует большое число масс-спектральных пиков, равномерно распределенных по шкале m/z . Следует подчеркнуть, что пик молекулярного иона присутствует и является основным. Во-вторых, наличие одного или нескольких атомов хлора в составе молекулярного и ряда фрагментных ионов приводит к появлению дополнительных изотопных пиков, что увеличивает достоверность оценки параметра t_{overhead} .

Первый этап работы был посвящен проверке нашей гипотезы о наличии линейной зависимости между временем удерживания и значением m/z соответствующей масс-хроматограммы. Здесь и далее мы подразумеваем, что время удерживания определено с точностью, превосходящей шаг дискретизации. На рис. 5 представлено несколько таких зависимостей. При интерполяции параболой по трем точкам число степеней свободы равно нулю, и погрешность определения положения максимума оказывается ожидаемо высокой, что не позволяет выявить какую-либо зависимость (рис. 5а). При аппроксимации пика параболой по пяти точкам число степеней свободы равно двум. На соответствующем графике (рис. 5б) уже прослеживается зависимость между положением вершины пика и значением m/z соответствующей масс-хроматограммы. Наилучшие результаты были получены, когда во внимание принимали все точки хроматографического пика (и проводили аппроксимацию функцией Гаусса или ЭМГ). В качестве примера на рис. 5в представлена зависимость, полученная при аппроксимации пика функцией ЭМГ (с фиксированными значениями σ и τ). В этом случае уже можно говорить о наличии линейной зависимости. Следует отметить, что при увеличении концентрации аналита зависимость проявляется более явно (так, на рис. 5г представлены результаты, полученные при анализе раствора с концентрацией 10 нг·мкл⁻¹).

Следующий этап работы был посвящен выбору наилучшего способа расчета положения вершины хроматографического пика. В настоящей работе определение параметров функции развертки основано на аппроксимации зависимостей, представленных на рис. 5, прямыми линиями. В связи с этим, чем меньше разброс точек относительно аппроксимирующей прямой, тем достовернее оценка параметра t_{overhead} . Разброс точек можно оценить

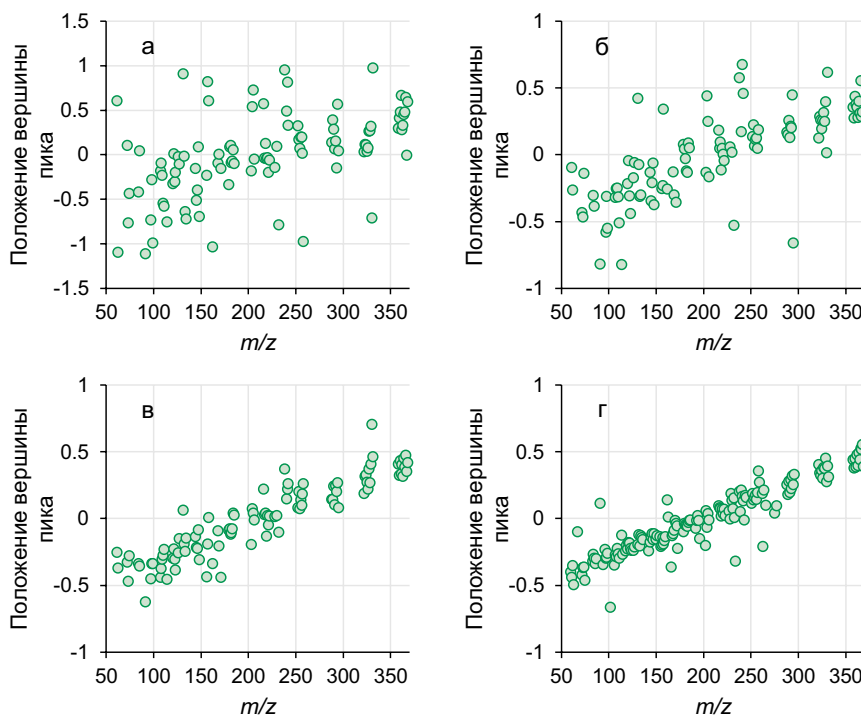


Рис. 5. Зависимость времени удерживания (рассчитанного с точностью, превосходящей шаг дискретизации) от соответствующего значения m/z . Положение вершины хроматографического пика было вычислено в результате интерполяции параболы (а), аппроксимации параболы (б) или аппроксимации экспоненциально модифицированной функцией Гаусса (в, г). Концентрация гексахлорбифенила составляла 1 нг·мкл⁻¹ (а, б, в) или 10 нг·мкл⁻¹ (г)

Fig. 5. Dependence of the retention time (calculated with an accuracy greater than one sampling interval) on the corresponding m/z value. Position of the apex is calculated using parabola interpolation (a), parabola approximation (б) or approximation by exponentially modified Gaussian function (в, г). The concentration of hexachlorobiphenyl is equal to 1 ng ul⁻¹ (а, б, в) or 10 ng ul⁻¹ (г)

количественно, рассчитав, например, остаточное стандартное отклонение:

$$s_e = \sqrt{\frac{RSS}{n-2}} = \sqrt{\frac{\sum e_i^2}{n-2}},$$

где RSS – остаточная сумма квадратов; e_i – разность между фактическим и расчетным значением зависимой переменной (наблюдаемая в i -ой точке); n – число аппроксимируемых точек.

Как видно из рис. 6, во всех случаях наилучшие результаты были получены при аппроксимации пика функцией Гаусса или функцией ЭМГ с фиксированными значениями σ и t . Следует отметить, что в нашем исследовании полихлорированные бифенилы элюировались из хроматографической колонки в виде симметричных пиков. Однако достаточно часто пики в газовой хроматографии имеют размытый задний фронт, что делает необоснованным применение функции Гаусса для их аппроксимации. Чтобы сохранить универсальность подхода, предлагаемого нами для определения параметров функции развертки, мы ограничились рассмотрением только функции ЭМГ (с фиксированными значениями σ и t).

Можно ожидать, что случайная погрешность времени удерживания зависит от отношения сигнал/шум соответствующего пика. В настоящей работе

эта зависимость была определена экспериментально (поскольку теоретический расчет представляет собой весьма сложную задачу). Для этой цели все значения m/z , присутствующие в масс-спектре рассматриваемого соединения, были разделены на группы в зависимости от величины отношения сигнал/шум соответствующего хроматографического пика ($s/n \leq 15$, $15 < s/n \leq 25$, $25 < s/n \leq 50$, $50 < s/n \leq 100$, $100 < s/n \leq 300$, $300 < s/n \leq 1000$, $1000 < s/n$). Для каждой группы рассчитывали медианное значение параметра s/n , а также стандартное отклонение времени удерживания, используя следующую формулу:

$$s_{Rt} = \sqrt{\frac{\sum (Rt_{mz} - (k \cdot mz + b))^2}{n-1}},$$

где суммирование проводили по всем значениям m/z , принадлежащим рассматриваемой группе; Rt_{mz} – время удерживания соединения на масс-хроматограмме, построенной по определенному значению m/z ; mz – значение m/z ; k и b – коэффициенты аппроксимирующей прямой, определяющей взаимосвязь между положением максимума хроматографического пика и значением m/z (рис. 4в); n – число значений m/z , входящих в рассматриваемую группу.

На рис. 7 представлена зависимость стандартного отклонения времени удерживания от отношения сигнал/шум соответствующего хроматографического

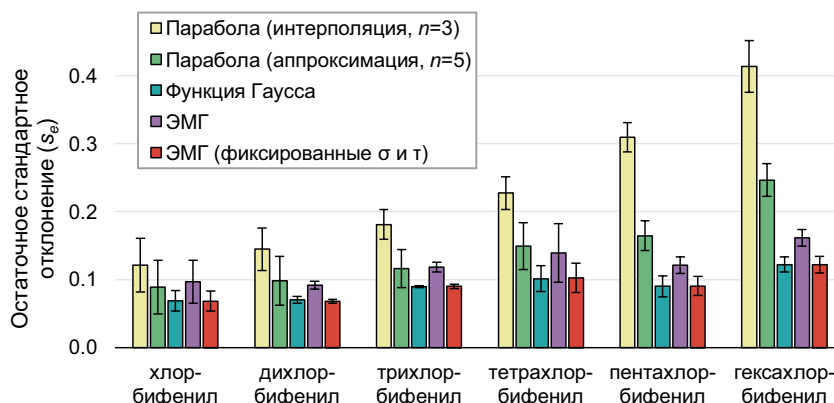


Рис. 6. Остаточное стандартное отклонение линейной регрессии (построенной для зависимости положения вершины хроматографического пика от соответствующего значения m/z). Вертикальные отрезки представляют собой стандартные отклонения ($n = 3$)

Fig. 6. Residual standard deviation of the linear regression (calculated between the position of the chromatographic peak apex and corresponding m/z value). Error bars are standard deviations ($n = 3$)

ского пика. На графике одновременно отображены результаты, полученные для всех модельных соединений и всех параллельных экспериментов (число которых было равно трем). Представленная зависимость хорошо аппроксимируется степенной функцией, которая в билогарифмических координатах трансформируется в прямую линию (рис. 7).

Уравнение степенной функции, представленное на рис. 7, можно использовать для оценки погрешности времени удерживания, зная величину отношения сигнал/шум соответствующего хроматографического пика. В связи с этим для аппроксимации зависимостей, представленных на рис. 5в и 5г, можно использовать взвешенный метод наименьших квадратов [13], снижая при этом вес точек, ординаты которых определены с относительно большой погрешностью:

$$\mathbf{B} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y},$$

где \mathbf{B} – вектор коэффициентов уравнения линейной регрессии; \mathbf{X} – матрица независимых переменных (константы, равные единице, и значения m/z); \mathbf{Y} – вектор зависимых переменных (времена удерживания, определенные с точностью, превосходящей шаг дискретизации); \mathbf{W} – диагональная матрица весов (диагональный элемент которой w_{ii} обратно пропорционален квадрату стандартного отклонения y_i). Следует подчеркнуть, что использование взвешенного метода наименьших квадратов позволило в 2-3 раза повысить воспроизводимость оценки параметра t_{overhead} .

Значения параметра t_{overhead} , рассчитанные при использовании предлагаемого нами подхода, представлены на рис. 8. Даже невооруженным глазом видно, что наихудшая воспроизводимость наблюдается в случае раствора, обладающего наименьшей концентрацией. Такой результат ожидаем, поскольку при концентрации 0.32 нг мкл^{-1} только для 7-10 значений m/z отношение сигнал/шум (на соответствующей масс-хроматограмме) превышает 100. При увеличении концентрации модельного раствора

до 1 нг мкл^{-1} число таких значений m/z возрастает в среднем в 2.5 раза (а общее число масс-спектральных пиков возрастает в 2 раза), что увеличивает достоверность оценки параметра t_{overhead} .

Можно предположить, что достоверность оценки параметра t_{overhead} определяется не только количеством анализа, введенного в прибор, но и тем, какую часть диапазона сканирования покрывают пики, присутствующие в масс-спектре. В настоящей работе масс-спектральные пики покрывали от 44 (в случае хлорбифенила) до 99 (в случае гексахлорбифенила) процентов диапазона сканирования. К сожалению, на данном этапе мы не можем ни подтвердить, ни опровергнуть высказанное предположение, поскольку нам доступен относительно небольшой объем экспериментальных данных.

Истинное значение параметра t_{overhead} не отображается в программном обеспечении ChemStation (Agilent), управляющем масс-спектрометром

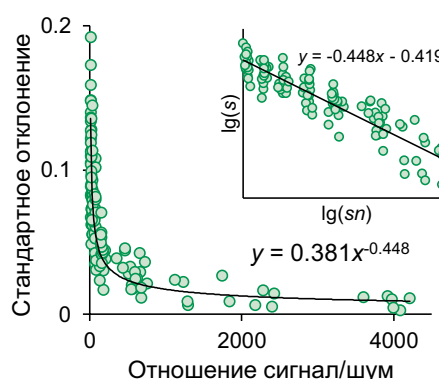


Рис. 7. Зависимость стандартного отклонения времени удерживания (рассчитанного с точностью, превосходящей шаг дискретизации) от отношения сигнал/шум. На вставке представлен график, построенный в билогарифмических координатах

Fig. 7. Dependence of the standard deviation of the retention time (calculated with an accuracy greater than one sampling interval) on the signal to noise ratio. In the insert, log-log plot is shown

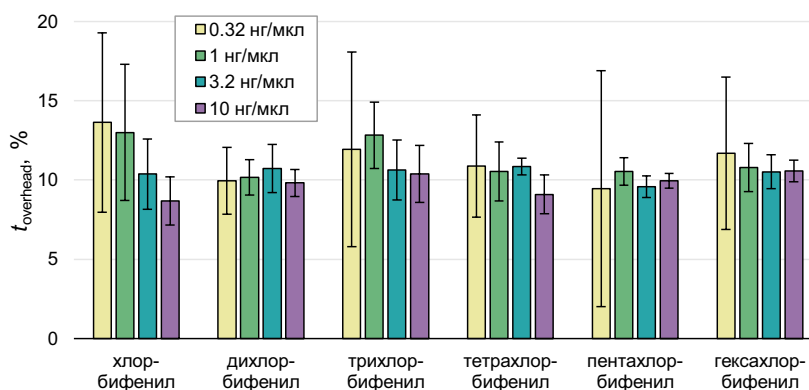


Рис. 8. Значения параметра t_{overhead} , рассчитанные при использовании предлагаемого нами подхода. Параметр t_{overhead} выражен в процентах от времени одного скана. Вертикальные отрезки представляют собой стандартные отклонения ($n = 3$)

Fig. 8. Overhead time calculated using the proposed approach. The overhead time is expressed as a percentage of the total scan time. Error bars are standard deviations ($n = 3$)

Agilent 5973. Тем не менее, в качестве истинного значения параметра t_{overhead} можно использовать среднее значение, рассчитанное для растворов концентрацией 3.2 и 10 нг мкл⁻¹ и равное (10.1 ± 0.4) % от времени одного скана ($n = 36$, $P = 0.95$). Согласно критерию Пирсона, значения параметра t_{overhead} распределены нормально ($n = 36$, $P > 0.49$). Следует отметить, что единичные оценки параметра t_{overhead} изменялись в достаточно широких пределах: от 7 до 13 % от времени одного скана. В связи с этим при использовании предлагаемого нами подхода в реальной практике целесообразно проводить усреднение параметра t_{overhead} . В качестве примера в табл. представлены оценки параметра t_{overhead} , рассчитанные при использовании всех шести соединений, присутствующих в модельном растворе. Как видно из представленных данных, истинное значение во всех случаях попадает в 95 % доверительный интервал (который в среднем составляет ± 2.3 % от времени одного скана для рас-

твора концентрацией 1 нг мкл⁻¹ и ± 1.3 % от времени одного скана для растворов концентрацией 3.2 и 10 нг мкл⁻¹).

При разложении фрагмента исходных (искаженных) ГХ/МС данных, содержащих хроматографический пик гексахлорбифенила (10 нг мкл⁻¹), методом главных компонент остаточная дисперсия первой главной компоненты составляет в среднем 0.6 %. При разложении фрагмента ГХ/МС данных, исправленных при использовании алгоритма, не учитывающего параметр t_{overhead} , остаточная дисперсия составляет 0.09 %. Учет параметра t_{overhead} позволяет дополнительно уменьшить остаточную дисперсию первой главной компоненты в полтора раза, до 0.06 %. Проведенная коррекция значительно повышает линейность данных и способна улучшить результаты работы алгоритмов автоматического поиска пиков компонентов и выделения «чистых» масс-спектров.

Заключение

В работе предложен математический подход, предназначенный для определения параметров функции развертки квадрупольного масс-спектрометра непосредственно из исходных ГХ/МС данных, содержащих отдельно стоящий хроматографический пик. Вычисления проводят в несколько этапов. Вначале на всех масс-хроматограммах, содержащих пик аналита, определяют положение максимума хроматографического пика, проводя аппроксимацию данных экспоненциально модифицированной функцией Гаусса. Затем при использовании взвешенного метода наименьших квадратов строят линейную зависимость между временем удерживания (рассчитанным с точностью, превышающей шаг дискретизации) и значением m/z . Знак коэффициента наклона этой прямой определяет направление сканирования, а абсолютное значение используется для вычисления параметра t_{overhead} . Показано, что воспроизводимость рассчитанных значений

Таблица

Средние значения и доверительные интервалы ($n = 6$, $P = 0.95$) параметра t_{overhead} , рассчитанные для 9 независимых экспериментов. Усреднение проводили по всем шести соединениям, присутствующим в модельной смеси. Параметр t_{overhead} выражен в процентах от времени одного скана

Table

Average values and confidence intervals ($n = 6$, $P = 0.95$) of the overhead time calculated for nine independent experiments. Data is averaged over all six compounds present in the model mixture. The overhead time is expressed as a percentage of the total scan time

Номер параллельного опыта	Концентрация модельного раствора		
	1 нг/мл	3.2 нг/мл	10 нг/мл
1	12.1 ± 3.3	10.3 ± 0.8	9.5 ± 0.7
2	10.8 ± 2.2	10.1 ± 1.7	10.0 ± 1.1
3	11.0 ± 1.3	11.0 ± 1.5	9.7 ± 1.9

параметра t_{overhead} главным образом зависит от количества аналита, введенного в прибор. Это объясняется тем, что уменьшение отношения сигнал/шум хроматографического пика в 10 раз приводит к увеличению погрешности определения положения его максимума в 2.8 раза. При анализе растворов модельных соединений концентрацией 3.2 и 10 нг·мкл⁻¹ относительное стандартное отклонение параметра t_{overhead} составило порядка 1.2 % от времени одного скана, для раствора концентрацией 1 нг·мкл⁻¹ — около 2.2 % от времени одного скана.

Благодарности

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 16-33-60169 мол_а_дк.

Acknowledgements

The presented study was funded by Russian Foundation for Basic Research, according to the research project No. 16-33-60169 mol_a_dk.

Литература

- Сравнительные возможности масс-анализаторов разных типов в решении аналитических задач масс-спектрометрическими методами / Л.Н. Галль [и др.] // Масс-спектрометрия. 2008. Т. 5, № 4. С. 295-300.
- Самохин А.С., Гусев А.Е., Каламбет Ю.А. Математический подход к устранению искажений масс-спектров, зарегистрированных на склонах хроматографических пиков // Аналитика и контроль. 2017. Т. 21, № 2. С. 144-152.
- Сысоев А.А., Чупахин М.С. Введение в масс-спектрометрию. М.: Атомиздат, 1977. 304 с.
- Smith R. *Understanding mass spectra: a basic approach*. New York: Wiley, 2004. 372 p.
- Watson T., Sparkman D. *Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation*. Chichester: Wiley, 2008. 862 p.
- King M.D., King G.S. A numerical method for extracting mass spectra from gas chromatography mass spectrometry data arrays // *Anal. Chem.* 1985. V. 57, № 6. P. 1049-1056.
- Pool W.G., de Leeuw J.W., van de Graaf B. A rapid routine to correct for skewing in gas chromatography/mass spectrometry // *J. Mass Spectrom.* 1996. V. 31, № 2. P. 213-215.
- Stein S.E. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data // *J. Am. Soc. Mass Spectrom.* 1999. V. 10, № 8. P. 770-781.
- Fraga C.G. Chemometric approach for the resolution and quantification of unresolved peaks in gas chromatography-selected-ion mass spectrometry data // *J. Chromatogr. A*. 2003. V. 1019, № 1-2. P. 31-42.
- Foley J.P., Dorsey J.G. A review of the exponentially modified gaussian (EMG) function: evaluation and subsequent calculation of universal data // *J. Chromatogr. Sci.* 1984. V. 22, № 1. P. 40-46.
- Reconstruction of chromatographic peaks using the exponentially modified Gaussian function / Y. Kalambet [et al.] // *J. Chemometr.* 2011. V. 25, № 7. P. 352-356.
- Delley R. Series for the exponentially modified Gaussian peak shape // *Anal. Chem.* 1985. V. 57, № 1. P. 388.
- Каламбет Ю.А., Мальцев С.А., Козьмин Ю.П. Доверительные интервалы метода взвешенных наименьших квадратов и стратегия градуировки // Заводская лаборатория. Диагностика материалов. 2015. Т. 81, № 1. С. 69-76.

References

- Gall L.N., Bezenov A.N., Kouzmin A.G., Gall N.R. [Comparative possibilities of various type mass-analyzers for analytical problems of mass-spectrometry]. *Mass-spektrometriia* [Mass spectrometry], 2008, vol. 5, no. 4, pp. 295-300 (in Russian).
- Samokhin A.S., Gusev A.E., Kalambet Y.A. [Mathematical approach to eliminating the distortions of mass spectra recorded on the slopes of chromatographic peaks]. *Analitika i Kontrol'* [Analytics and control], 2017, vol. 21, no. 2, pp. 144-152 (in Russian). DOI: 10.15826/analitika.2017.21.2.007
- Sysoev A.A., Chupakhin M.S. *Vvedenie v mass-spektrometriiu* [Introduction to mass spectrometry]. Moscow, Atomizdat, 1977. 304 p. (in Russian).
- Smith R. *Understanding mass spectra: a basic approach*. New York, Wiley, 2004. 372 p. DOI: 10.1002/0471479357
- Watson T., Sparkman D. *Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation*. Chichester, Wiley, 2008. 862 p. DOI: 10.1002/9780470516898
- King M.D., King G.S. A numerical method for extracting mass spectra from gas chromatography mass spectrometry data arrays. *Anal. Chem.*, 1985, vol. 57, no. 6, pp. 1049-1056. DOI: 10.1021/ac00283a020
- Pool W.G., de Leeuw J.W., van de Graaf B. A rapid routine to correct for skewing in gas chromatography/mass spectrometry. *J. Mass Spectrom.*, 1996, vol. 31, no. 2, pp. 213-215. DOI: 10.1002/(SICI)1096-9888(199602)31:2<213::AID-JMS284>3.0.CO;2-6
- Stein S.E. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *J. Am. Soc. Mass Spectrom.*, 1999, vol. 10, no. 8, pp. 770-781. DOI: 10.1016/S1044-0305(99)00047-1
- Fraga C.G. Chemometric approach for the resolution and quantification of unresolved peaks in gas chromatography-selected-ion mass spectrometry data. *J. Chromatogr. A*, 2003, vol. 1019, no. 1-2, pp. 31-42. DOI: 10.1016/S0021-9673(03)01329-3
- Foley J.P., Dorsey J.G. A review of the exponentially modified gaussian (EMG) function: evaluation and subsequent calculation of universal data. *J. Chromatogr. Sci.*, 1984, vol. 22, no. 1, pp. 40-46. DOI: 10.1093/chromsci/22.1.40
- Kalambet Y., Kozmin Y., Mikhailova K., Nagaev I., Tikhonov P. Reconstruction of chromatographic peaks using the exponentially modified Gaussian function. *J. Chemometr.*, 2011, vol. 25, no. 7, pp. 352-356. DOI: 10.1002/cem.1343
- Delley R. Series for the exponentially modified Gaussian peak shape. *Anal. Chem.*, 1985, vol. 57, no. 1, pp. 388. DOI: 10.1021/ac00279a094
- Kalambet, Y.A., Maltsev, S.A., Kozmin, Y.P. [Confidence intervals for weighted least squares technique and calibration strategy]. *Zavodskaya laboratoriya. Diagnostika materialov* [Industrial laboratory. Diagnostics of Materials], vol. 81, no. 1, pp. 69-76 (in Russian).